DOCUMENT RESUME

ED 400 167 SE 058 457

AUTHOR Wilson, Linda Dager
TITLE Documenting Observations of Students in Mathematics: A Case Study.
INSTITUTION National Center for Research in Mathematical Sciences Education, Madison, WI.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 10 Apr 96
NOTE 42p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Algebra; Case Studies; *Classroom Observation Techniques; Educational Research; *Grading; Informal Assessment; Institutional Cooperation; Intermediate Grades; Junior High Schools; *Mathematics Instruction; Participant Observation; Portfolio Assessment; *Student Evaluation

ABSTRACT

This paper is the report of a year-long study of a mathematics teacher's attempts to document her observations of students. One goal of the project is to determine whether it is feasible for a teacher to keep a record of informal assessments such as observations and interviews. Another goal is to determine how useful such information could be to the teacher. Two classes of Algebra I students (N=59) were part of this study which was a collaborative effort involving field notes, videotaping, survey results, audiotaped interviews, teacher notes and records of student progress. Background information is included about the early development of the project, the logistics of documentation methods, and key instructional decisions made by the teacher throughout the project. Important findings of this study highlight the importance of being clear about what is being assessed and the links between that and instruction. The findings also provide a picture of the challenges, from a teachers perspective, of aggregating observation and interview data with other records of student progress. Contains 13 references. (DDR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 400 167

Documenting Observations of Students in Mathematics: A Case Study

Linda Dager Wilson
Department of Educational Development
University of Delaware
Newark, DE 19716
ldwilson@udel.edu

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

L.D. Wilson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This research was made possible through funding from the Office of Educational Research and Improvement, National Center for Research in Mathematical Sciences Education.

BEST COPY AVAILABLE

SE059457

Documenting Observations of Students in Mathematics: A Case Study

In this paper I report on a year-long study of a mathematics teacher's attempts to document her observations of students as they worked in her class. The project was a collaborative effort to investigate whether or not it is feasible for a classroom teacher to keep a record of the kinds of assessment that are typically done informally in classrooms, namely, observations and interviews. Further, our goal was to determine how useful such information could be to the teacher for various assessment purposes, such as instructional decision making, monitoring student progress, and assigning grades.

Rationale

Assessment as evidence-gathering

The *Assessment Standards for School Mathematics* (National Council of Teachers of Mathematics, 1995) defines assessment as "gathering evidence about a student's knowledge of, ability to use, and disposition towards, mathematics and of making inferences from that evidence for a variety of purposes" (p.3). The conceptualization of assessment used in this study is based on this definition. The process of assessment involves two components, namely, evidence-gathering and inference-making. Teachers collect data about students and then make decisions. It is an inferential process that must be based on data.

Implicit in this conception of mathematical assessment is the essential centrality of the teacher's role. In school settings it is, after all, the classroom teacher who

presumably knows the students best. The teacher has the best vantage point from which to gather multiple pieces of evidence and then to aggregate that data and infer what the child knows about mathematics, how well the child can use mathematics, and the child's disposition toward mathematics. Many recent assessment documents (e.g., NCTM, 1995; Stenmark, 1991) recommend that mathematics teachers make use of a wide variety of strategies for assessing students, including observations, interviews, portfolios, journals, and the like. Much of the evidence gathered from these sources is not easily quantifiable, suggesting a more interpretive approach. This view of assessment casts the teacher in a role not unlike that of an educational researcher, gathering data and making inferences based on the evidence at hand.

The evidence about student learning that is available to teachers in classrooms comes in a variety of forms, but there are three primary sources: observations, interviews, and the work that students produce. Teachers have opportunities to watch students at work, to observe the strategies they use to solve problems, or the behaviors they display in class. They can listen to the explanations students give, and ask probing questions to discern the depth of their understanding. And they can examine the products of learning: written work, projects, reports and the like. There is evidence that, while teachers acknowledge that students most readily reveal their mathematical knowledge orally and through observable actions, teachers are reluctant to rely on this type of data when the resulting decisions are open to scrutiny, such as determining grades or promotion. They are more likely to use such evidence for less scrutinized purposes, such as instructional and curricular decision-making (Dorr-Bremme & Herman,

1986; Watson, 1995).

There are varying interpretations of how the role of teacher as assessor/researcher might be played out in the classroom. Mellin-Olsen (1993) describes the teacher-as-researcher activities as "practical hermeneutics," suggesting that teachers engage in a kind of discourse analysis of data from interviews with students. The methods he offers are essentially the methods of qualitative data analysis, interpreting the text to determine the meaning inherent in them. Ginsburg suggests clinical interviews as the primary sources of data for "thinking assessment" (1993). Freudenthal argued for more naturalistic observations of students' thinking (cited in van den Heuvel-Panhuizen, 1996). Regardless of the techniques advocated, there is no question that the use of observations and interviews has been essential to research on students' thinking in mathematics, and that some version of these techniques are currently being advocated as strategies for classroom assessment. Rather than relying solely on quantitative measures, teachers are being asked to make use of more and more qualitative data in mathematical assessment.

Links to pedagogy, curriculum, and learning

From a traditional measurement perspective, assessment tends to be viewed as external to, or isolated from, the activities of teaching and learning (Graue, 1993). Yet the act of gathering evidence and making inferences about student learning is undeniably connected to the other decisions and actions in which teachers are engaged. For example, the evidence a teacher gathers, informally and formally, is used to make

instructional decisions about what to teach and also how to teach. Evidence about the effectiveness of a lesson, or of an entire program, is used to make curricular decisions. In addition, evidence about students' responses to a curriculum is used to monitor student progress toward instructional goals. In these ways, assessment informs the linkages between and among teaching, learning, and the curriculum.

The links among assessment, teaching, and learning are central to Graue's (1993) notion of "instructional assessment." She conceives of assessment as an "ongoing interpretation of student learning within instruction" (p. 289). From this point of view, teachers and learners are engaged in a recursive and collaborative process that continually informs the pedagogical and curricular decisions the teacher makes. Webb (1992) describes the features of instructional assessment as a teacher who sets expectations for student learning, gathers and interprets information about student learning, and then makes informed decisions throughout instruction. Van den Heuvel-Panhuizen (1996) calls this type of assessment "didactical." In didactical assessment the purposes, content, procedures, and tools for assessment are all linked to instruction and learning. In all of these interpretations, assessment is inseparable from teaching and learning.

Classroom assessment as evidence-gathering and inference-making is a central node in the web of links among teaching, learning, and the curriculum. The connections among these elements suggests that the use of alternative and more varied forms of evidence about student learning might have consequences for these other nodes in the web of classroom activity as well.

In their model of "thinking assessment" Ginsburg et al (1993) conjecture that the use of observations and interviews sends new messages to students about what is valued, namely, thinking, metacognition, and the importance of communicating mathematical ideas in a variety of forms. They also claim that thinking assessment demands that teachers model thinking for children and that they reflect on their own practice. Thus the curriculum shifts to a focus on mathematical thinking. Further, such activities can alter a teacher's sense of self, moving from being an authority figure to a fallible human being.

The conception of assessment as evidence-gathering and inference-making requires that teachers engage in more of the activities of research, gathering data from a variety of sources and testing hypotheses based on that evidence for different purposes. The use of observations and interviews would have to be central to this view. But how feasible is it for the typical classroom teacher to gather this kind of evidence, and of what practical use is this information? Furthermore, if it is feasible, what effects might these forms of assessment have on the teacher or the students? These are the questions that formed the foundation of this study.

Methods

An eighth grade mathematics teacher (referred to in this paper as Ms. Vance) agreed to collaborate with me in an investigation of the feasibility of documenting observations of students. Ms. Vance experimented with different strategies for recording what she observed in her two Algebra 1 classes. My role in this year-long study was to offer suggestions when solicited, and otherwise to systematically observe and analyze the

experiment as it unfolded. My research goals were to not only describe the strategies that Ms. Vance used for observational assessment and the use to which she put the resulting information, but to look for possible effects of these activities on her teaching and on students.

I observed both Algebra I classes approximately once a week throughout the school year, keeping fieldnotes. At the time of each visit I also met with Ms. Vance and audiotaped our conversations. Six of the classes were videotaped. Ms. Vance made audiotapes of interviews with her students, and these were part of my data, as well as notes she kept and records she made of student progress. In addition I administered a survey to the students at the beginning and end of the school year and audiotaped interviews with selected students. Other sources of data include copies of tasks done with students and data from a gender study that Ms. Vance was involved in.

The Setting

The study was conducted with two classes of eighth grade Algebra 1 students in a public middle school in the Mid-Atlantic. There were thirty students in one class and twenty-nine in the other. Thirty-nine students were male and twenty female, eight were nonwhite and fifty-one white in the two classes combined. The school operated on a three-day rotating schedule, so these classes met for about 65 minutes on two of the three days in the rotation. Every school day began with a forty-minute planning period for teachers, followed by two classes, then a half-hour lunch, another forty-minute planning period, and two more classes. Besides the two classes of Algebra 1, Ms. Vance taught one general mathematics class and one class of reading. She would often meet

with the other teachers in her "reading team" during the planning periods to plan the reading classes.

Ms. Vance has twenty years' experience as a mathematics teacher, both at secondary and middle school levels. She holds certification at the middle and secondary levels in mathematics. She characterized herself as having been a traditional teacher until approximately three years ago, when she began to work with a state-wide NSF-funded Teacher Enhancement Project and with the New Standards project. One of the primary goals of the Teacher Enhancement Project was increasing teachers' pedagogical content knowledge (Schulman, 1986) in mathematics. For that project Ms. Vance was named a resource teacher for her school, meaning she was designated as a leader for the other mathematics teachers in her building. She also conducted several content-specific workshops for other teachers in the project. As part of the New Standards project Ms. Vance piloted portfolio materials with the classes I observed.

The pedagogical approach in the Algebra I classes could be characterized by an emphasis on problem solving and on explaining solution strategies. As we shall see, Ms. Vance was quite clear that her primary instructional goal was problem solving, and this was apparent in her choice of curricular materials. She rarely used a textbook, but made wide use of materials from a variety of sources, including New Standards, the PACKETS program from Educational Testing Service, and Math in Context (a contextualized middle grades curriculum). The focus in class was on finding ways to solve problems, often in collaboration with partners, and in sharing solution strategies. The latter was usually accomplished through whole class discussions. "How did you get that?" and "Did

anyone do it another way?" were questions that permeated every lesson that I observed.

A typical class day might begin with a whole class game of mental math. Ms. Vance would bring a deck of index cards, on which had been printed something like, "I have 53. Who has that times 2?". When that card was read aloud, the student holding the card with "I have 106. Who has 90 less?" would be expected to read his or her card aloud. Ms. Vance would pass out the cards, one to each student, and then begin the game with her own card. Play would proceed around the room until the final answer came back to her. One of the students would time the game, and the goal was to minimize the total time. She had sets of cards for operations on whole numbers, integers, fractions and decimals. A typical game would last 4 or 5 minutes.

Ms. Vance would often introduce a new topic, such as linear functions, with a task to do in pairs. An example task had information about two video stores, one that charges a membership fee and has a low rental rate and a second one that has no initial fee but a higher rental rate. The problem was to figure out which store has the better deal. Students were expected to work in pairs to solve the task, but turn in separate solutions. The worksheet gave some guidance on how to approach the problem (such as making a chart and then a graph), followed by questions that required students to explain why their solution is correct. A task might take two or three days to complete. As the students worked through a task, Ms. Vance circulated around the room, encouraging students to think through the problem. She might interrupt the class for a whole group discussion if she became aware of a widespread misconception or a need for some background information. Once everyone had worked through the task, she would

typically conduct a whole class discussion, asking for students to explain their answers and compare strategies.

The core of the curriculum centered around about fifteen major tasks or projects. In one long-term project students were instructed to design a cereal box that minimized surface area and maximized volume. Another task was to discover and generalize patterns in the "Twelve Days of Christmas" carol. Formal presentations were sometimes videotaped and then reviewed again with the class. Students were required to retain their work on such tasks and put together a portfolio at the end of the year, writing about why they chose certain tasks or projects for inclusion, and critiquing each others' work.

In between these extended tasks, Ms. Vance would conduct classes of direct instruction on related concepts and skills, such as slope or the graphs of quadratic functions. A common assignment at the end of such a class was to write about the mathematics they had learned that day.

The classroom was large and bright, the back wall having tall windows that looked out on the front parking lot and the street beyond. Hanging from the ceiling and tacked to the walls were examples of student projects, such as polyhedra. Above the blackboard in the front were various mathematical posters, mostly with visual puzzles to solve. An overhead projector, with a panel for use with the graphing calculator, was positioned at the front of the room. A printed approximation of pi filled the wall above the student lockers that lined one wall. Beneath the windows were shelves equipped with a wide assortment of manipulatives and other tools. A set of thirty TI-81 calculators hung on

one wall. Student desks faced the blackboard in rows, while Ms. Vance's desk was positioned perpendicular to the students along one side. It was usually stacked high with papers and memos, and was used primarily for storage and not as a working space.

Early Trials: How and What and Who

During the first two months of school Ms. Vance tried an assortment of methods for recording her observations of students. The first method was to simply write some notes to herself at the end of each class. This she quickly abandoned as unfeasible, when she found that the three minutes' passing time between classes did not allow for such a high-concentration activity. She then tried to write notes at the end of the school day, but she found that most of the specific details of what she had observed were blurred at best or had disappeared entirely from memory by the end of the school day. (Examples of these notes will be examined in a subsequent section.) It became apparent that she would need to find a method that could be used in class at the time the actions were occurring.

I suggested a tape recorder when she told me that she was just not comfortable writing her observations down. She explained that writing took too much time and she tended to lose track of the slips of paper. With the tape recorder, she again found that waiting until the end of class or the end of the day was not effective, so she began to bring it to class. During the first few days with the tape recorder, a phenomenon occurred that we referred to as the "E.F. Hutton effect." While students were busy working in pairs, Ms. Vance would periodically pause to record what she was observing. Each time she turned on the tape recorder, the normal buzz of students at work would

stop and every ear would strain to hear what she was saying. The students were obviously curious and perhaps a bit suspicious of what she might be saying about them. Ms. Vance explained what she was doing and let them listen to what she had said, which seemed to quell any anxieties. A few days later she decided to try letting students have their turn with the tape recorder. As she circulated among the pairs, she would periodically ask them questions about their strategies or their conclusions and record these interviews. As she moved from that group to another, she would add her own commentary of what she had observed. This method proved to be popular with the students, who would often clamor to have their moment on tape. For awhile she did not abandon the written notes, but used a combination of notes and recordings, depending on her mood. Eventually she preferred the tape recorder exclusively, arguing that having it in hand was a good reminder to use it.

While the logistics of methods of documentation arose early, a more substantial issue became apparent when Ms. Vance reflected back on the early notes and recordings she had made. That was the issue of <u>what</u> to document. What follows are two excerpts from written notes and tape recordings done early in the year. An analysis of these will show that early attempts to document were largely unfocused and not particularly useful for assessment purposes.

> *(Notes written after a class in October, when students were working on a task to design four-cube houses to certain given specifications)*
>
> Mark & Jonathan
> Good statement about each home. Mark good movement. No need to go over. Include
>
> Jessica & Jessica

Good project scale report

Wendy & Stacy
Weak

Ethan & Robert
Exterior?

Mestafa
Reason for recommendation. Attractive brochure--need to spice it up.

(Recording made at end of class in late September, when students had been working on a spatial visualization task)

Remind myself never to do groups in the afternoon for the first time. Number one. Number two, much more work needs to be done in groups processes. I think they definitely have to have a much more challenging task before they can even attempt to do groups.
Who did I notice? Gary Moore had a difficult time in visualization. He had a difficult time building and needed my assistance plus the assistance of the people in his group. By the end of the period things had gotten a little bit easier for him. Malcolm--no not Malcolm--Russell could shade when we did the puzzle pieces too. He was shading cubes but he wasn't shading the entire face and he was confused on which faces to shade. Those were the two students that I especially took notice of -- Russell and Gary.

These two examples are typical of the kinds of documentation Ms. Vance made in the first three months of the school year. Some notes are too vague (e.g., Wendy and Stacy are "weak"), while others are task specific. The recorded comments refer to the whole class, such as the references to group process, while the written comments tend to focus on a few individuals. The most salient aspect of these early attempts is their lack of focus. It's as if the camera lens is continuously zooming in and then zooming out, with the result being a collection of information that is not very useful for any assessment purpose. Had she been clear that the purpose of the notes was for

instructional decision making, the notes might have remained specific to the task, and been useful for making more informed decisions about what to do tomorrow, next week, or next year. On the other hand, if the purpose was to monitor student progress, the information in its present form is useless without some sort of links from one task to the next.

Ms. Vance began to recognize this dilemma sometime in November, during an interview in which she was reflecting back on the set of written notes from the "Model Houses" project given above:

Ms.V: Yes and this is nothing.

LW: This early one is kind of--well some of it is about their report. Some of it is about--what *is* it about?

Ms.V: I can't read it. I mean I can *read* any of them for you: "Good statement about each house--made good movement--no need to go over--good--I don't know--good projects--scale report--they were weak" Just weak. That's all I put down.

LW: Uh huh.

Ms.V: "Extended" I guess. "Reason for recommendations"--he put them in, that means. If I wrote it, that means he did it. "Attractive brochure" and he needed to spice himself up when he did his presentation. Again, it was nothing. Nothing that I could say, "Oh, that means that they've had this." Nothing.

LW: So not only did you not like writing things down, but you haven't found this stuff very helpful since then, or have you?

Ms.V: No. To look back at what they did, that wasn't helpful at all to me.

By the end of two months of school, Ms. Vance was making progress on the logistical problems of what methods to use, but larger questions remained about what

BEST COPY AVAILABLE

she was looking for and who (whole class or individuals) she was targeting. As she attempted to organize the information she had gathered, an idea emerged that answered these questions and also began to solve the problem of how to use the information.

A Breakthrough

In mid-October, Ms. Vance and I had a conversation about the kinds of information she had been able to document so far. I pointed out to her that some of the notes and recordings she had made were substantively about aspects of problem solving. I asked her if it would be useful to think of it as evidence toward some broad instructional goals. She replied that "math power" was her major goal, and, yes, she did think that this kind of information might be used as evidence of progress toward developing "math power."

Here is how Ms. Vance later described this "breakthrough:"

> It also became evident that I needed to decide exactly what I was looking for. When I read the notes I had written about the students, I saw that they were haphazard and not connected to each other. They were not likely to help me assess my students. I also realized that some aspects of doing mathematics were more suited to other forms of assessment. Isolated skills, factual information and algorithmic procedures could easily be assessed by a paper-and-pencil test, but what I was attempting to document were the process goals of mathematics instruction: problem solving, communication, and critical thinking skills. These were the aspects of doing mathematics that I thought this form of assessment could help me with (Vincent & Wilson, 1996, p. 249).

By the end of the conversation we had explored the idea of developing some sort of "master rubric" that might specify exactly what her broad instructional goals were. At this point it was unclear exactly how she might use the rubric, but it seemed essential for her to articulate exactly what she wanted her students to accomplish during the year, in terms of processes such as problem solving or communication or group skills. Once that

task was accomplished we both hoped it would become clearer how this rubric would fit into the observations she had documented. Over the course of the next few weeks we discussed some of the general forms that this rubric might take.

It was not until the Thanksgiving break that Ms. Vance had the time to develop the rubric. To accomplish the task, she first studied several examples of published generalized mathematics rubrics, such as those from Vermont and Kentucky, and the New Standards project. Through her work with the Teacher Enhancement Project and New Standards, she had access to such materials. Building on those and choosing the aspects that fit her own vision of mathematical power, she came up with eight aspects of problem solving and then seven more aspects of group and individual work. She brought her first draft to the next meeting we had in December, and after some discussion she made some minor adjustments to it. The final product can be seen in Appendix A. From her own reports, developing this rubric was not difficult, but for her it was critical that she think it through herself, and not accept a "packaged" rubric as her own. This notion had implications for both assessment and instruction, as we shall see.

Another piece of the puzzle fell into place with the realization that documenting observations could not be done during every class period. As Ms. Vance discovered, there were many days when there was, simply, nothing to say. It could have been a day of predominantly direct instruction, when there was no opportunity to circulate around the room and observe individuals or groups at work. When she was leading a class discussion, her primary focus of attention was on the class and not on the individual learners. In other words, the class and not the students was the unit of her analysis.

A brief analysis (based on videotapes) of two "typical" classes, two days apart, illustrates this. On Day One, students had been working on the video store task the previous day, and Ms. Vance saw the need for some direct instruction on slopes and intercepts of linear equations. On Day Two, students were beginning a new task, the Twelve Days of Christmas, and working in pairs. Ms. Vance circulated around the room, interviewing students with tape recorder in hand. An analysis of the questions asked by Ms. Vance in each of those classes reveals that there were a total of 104 questions on Day One and 87 on Day Two, with generally the same breakdown of factual/procedural questions and higher order thinking questions (roughly 50% each). The difference lies in the number of questions asked of individuals about how they solved a problem. On the day of direct instruction Ms. Vance asked 6 such questions, while on the day of observational assessment she asked 29 "how" questions. Her intention on the first day was to move the whole class in its thinking about linear graphs, while on the second day she was interested more in facilitating each students' problem solving skills. From an assessment point of view, the opportunities for monitoring student progress were not present on the first day to the same extent that they were on the second.

But even on days when they were working on a project, if they had already mapped out their strategies and were simply carrying them out, the problem solving activity was limited to the execution of a solution strategy. For example, on one such day Ms. Vance told me, "They were all busy and they were all excited, but I didn't have much to say today about their progress in problem solving." Generally she found that the best days to use the tape recorder were on the days when students were engaged in

BEST COPY AVAILABLE

work on a longer task of some sort (one that might take one class period or a task that extended over several days), but were still in the early stages of planning a solution.

Having settled into a routine of when to record observations, a natural link between that and the master rubric was the next insight. Ms. Vance had the idea of creating a large chart on poster board that listed the students along one axis, and her master rubric across another axis. With the chart in front of her, she then listened to the tapes and recorded checks in boxes next to each students' name (see Figure 1). Check marks were coded to show the task they referred to. The chart became a record of each students' progress in problem solving over the course of the year.

Insert Figure 1 about here

By the end of the school year, Ms. Vance had been able to record her observations on the chart six times, corresponding to six of the extended tasks that her students had done. When I asked her about how she was able to find the time to listen to the tapes and record the checks, she assured me that the process went quickly. Her comment was that knowing the rubric so well, and internalizing it, made it easy to selectively listen to the tapes for the aspects of problem solving and group process that were in the rubric.

Reflecting back on the development of the rubric in early February, Ms. Vance observed:

> After I developed that rubric -- when was that? Right before Christmas? That helped to pull it all together for me. What I wanted to focus on. What was I

> trying to do with these kids? You know, what was going to be my mission? What's my end product going to look like? Then even breaking that down even further for the recording part of it, you know, how was I going to record all of this as they made progress? What was I going to call my levels of breaking that down? That to me was the real -- the "Ah hah!" moment. That made a big difference.

Clearly the rubric and the chart began to answer many of the questions of what to look for and who to observe.

Using the Data

Now that Ms. Vance had settled on a strategy for recording observations and interviews with students and for then tracking this information with her master rubric, of what use was all of this information? Conversations with her, as well as her own reflections (Vincent & Wilson, 1996) show that the primary use was for monitoring student progress towards her instructional goals, but there were secondary uses in making instructional decisions and for grading purposes.

Early in the school year, when she had not yet found a focus for her observations, many of the notes she wrote or recorded dealt with specific content or a particular task. For example, one day in a whole class discussion of a task involving counting the number of cubes in a pyramid from a two-dimensional drawing, a student named Mark had come up to the board to explain his solution. What he wrote can be seen in Figure 2.

Insert Figure 2 about here

At the end of that class, Ms. Vance recorded that she had not stopped the class to talk about Mark's irregular notation, but she reminded herself to go back and spend some

time on equations in the future. Even if she had not listened to this tape soon afterwards, the very act of reflecting back on that moment at the end of class helped to increase the chances that she would remember it and act on it later. In this way the documentation of specific content was useful for instructional decision making. However, she never organized these comments in any way, so that their usefulness was limited.

The primary use of the documented information was for monitoring student progress. This became clear once the method of charting students' progress on the chart was established. What Ms. Vance had developed was a means for gathering evidence about how well students were progressing towards her goals of problem solving, group skills, communication, and the more general "mathematical power." By the end of the year she could retrace the progress they had made from one task to another, or see general trends in their progress from one level to another. Because these same tasks formed the core of the students' portfolios, Ms. Vance decided that the information on the chart should be included in the portfolios. Together with the written work on each task and the reflective entry forms the students filled out on each task, the portfolio became a record of each students' progress in problem solving during the year. Ms. Vance gave students an overall score on their portfolio, and this took the place of a final exam for Algebra I. In this way, the documented evidence also served the purpose of evaluating student achievement. Her great frustration was in having to quantify the portfolios into single letter grades at the end of the year. Her school system did not allow for any kind of narrative reporting of student progress, so she was forced to

develop some arbitrary system for assigning number values to the portfolios, in order to obtain a numerical average.

Consequences of the Experiment

For Ms. Vance, participating in this experiment turned out to be, according to her, an "enlightening" experience. The consequences of trying to document observations of students were significant in terms of the impact on her teaching. Some of these positive consequences were a result of the process of identifying her instructional goals for the year.

Even before she formally articulated her goals in the "master rubric," Ms. Vance gave direct clues to her students about what she expected of them and what she valued. On the first day of class, as the students were discussing solutions to a puzzle problem involving clues about the number of eyes and legs of each animal, etc, she emphasized two aims: 1) the process is more important than the answer; and 2) they must learn to explain their solution both orally and in writing. Here are some excerpts from comments she made to the students in that class:

> More important than the answer is, how did you do it? What was your thinking?
>
> I'm going to ask you to share your solutions with the class, along with your strategy.
>
> Don't jump out with an answer. But how did you do it? We're all going to listen. See if you can understand by what he writes. Focus on the logic, not the answer.
>
> Put your name on [their written solutions] and pass it in. I want to see how well we explain ourselves. I'm not grading this. I want to know where to start in explaining our thinking.

These classroom excerpts illustrate many of the goals that Ms. Vance formalized in the

rubric. Problem solving, using appropriate strategies, is clearly a goal. She is emphasizing the solution process more than the final answer, and she places a high value on communication, i.e.,explaining solutions, listening to other explanations, and expressing solution processes in writing.

When she first tried to verbalize her instructional goals, Ms. Vance used the cryptic phrase, "math power." She spoke of "giving these children math power." Originating in the *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), the phrase has been used in many of the reform efforts to encapsulize goals for students. No doubt Ms. Vance had picked up this slogan from the reform efforts she was involved in (the NSF Teacher Enhancement project, the state curriculum framework commission, or possibly the New Standards project). What did Ms. Vance mean by math power? The rubric she developed indicates that problem solving, communicating mathematics, reasoning mathematically, and having a positive disposition about working alone or with others were the primary aspects. In addition to those aspects emphasized on the first day of school, the rubric includes social norms (working with others) and disposition. The social emphases are on working cooperatively, being an active member of a group, and considering the ideas of others. For individual learners the rubric stresses perseverance, risk-taking, and the ability to assess one's own work and revise accordingly.

In reflecting back on the school year, Ms. Vance felt strongly that the process of developing the rubric had been the most critical component in shaping her teaching. It gave her, she said, the opportunity to make clear (mostly to herself, and then to her students) what goals she had for her students. More than once she emphasized that, had

BEST COPY AVAILABLE

she simply adopted a ready-made rubric from the start and used it to record observations without going through the process herself, the effect on her teaching would not have been as positive as it was. Even while building on the work of others, it forced her to think hard about what she wanted for her students. She said that articulating these goals in turn helped her internalize them, which brought greater focus to her teaching.

The process of articulating her instructional goals had several consequences for instruction in the two Algebra 1 classes.

Knowing what to assess

The first consequence was on *what* assessment evidence Ms. Vance gathered from class. In an interview at the end of the year, when she was reflecting back on the effects of the year's experiment, Ms. Vance claimed:

> It helped me focus my questions that I want to ask -- the kind of reflective questions I would ask the students to gain even more insight into their thinking. I knew what kind of questions should come next.

Short of doing an in-depth discourse analysis of the classes, I do not have sufficient evidence about whether or not there were any changes in the questions she asked her students later in the school year. What I do have evidence for is the assessment record she made from classes early in the year and later, and the differences in what she attended to.

During one class in late September, Ms. Vance asked students to share their solution strategies for a spatial visualization problem they had begun in class the day before and finished for homework. Students had been given a diagram of a tower of cubes, with a pyramid shape cut out of it, and the task was to reconstruct the drawing on

isometric dotty paper and develop a strategy for counting the total number of cubes needed to build the tower. Ms. Vance listened while four students explained how they drew the figure on the dotty paper, and then she asked for solutions to the "how many cubes?" question. Gabe and Mark both offered answers, and each in turn was asked to explain his solution to the class. The class continued, with Ms. Vance explaining her strategy and then leading a discussion of how the pattern could be generalized.

> At the end of class, Ms. Vance recorded the following observations:
>
> Okay the second period Gabe Bowman's description of how he calculated 95 cubes was absolutely superior. He was clear; he was concise and used mathematical language and the whole class understood it. Also in the second period class I noticed that Mark Redding had a difficult time understanding what an equation was. Mark thought he had an equation and it truly was an equation. I'm going to have to make sure that I go back to that. What else I've noticed: in a general sense I've noticed that problem 4 and 5 was difficult for many, many students.

What is noticeable here about what Ms. Vance attended to in her recorded observation after the class is its focus on communication. Though during the class she had asked both boys many questions about their strategies and how they had solved the problem, the recorded evidence does not include this, but rather centers on the quality of the verbal and written explanations they gave.

In mid-December, after she had developed the master rubric and the chart for recording observations, Ms. Vance conducted a class in which students were working on a task in pairs. After a few moments of whole group discussion to introduce the Twelve Days of Christmas task, students worked in pairs while Ms. Vance circulated with her tape recorder and offered help, checked on progress, and interviewed students. What follows is an example of a typical exchange:

Ms. Vance: How's it going? Any problems?

Tracy: We're working on the pattern.

Ms. Vance: Say what you have in each column.

Wendy: This is the number of presents on each day, and this one is the total presents. On Day Two we have 3 presents here and 4 presents here.

Ms. Vance: Three presents on which day?

Later Ms. Vance played back this recording and placed checkmarks on the big chart according her master rubric. She made decisions about the levels of understanding, the students' ability to generalize, the level of their communication, and their ability to work together.

While the questions she asked of the students may not have been much different in September than they were in December, she was able to assess a broader range of learning attributes for each student because she had a more complete record of each students' thinking. Because of the generalized rubic, the data she collected was more focused. The evidence she gathered after the rubric was developed was organized and accessible. The process of making inferences from the data was based on a conceptual framework, which allowed her to make relatively quick and focused decisions about student learning.

Janice

Another consequence of observational assessment was evident in the case of one student, Janice. Early in the year Ms. Vance stated that one of her personal goals was to encourage the girls in her classes to participate and feel confident about their abilities in

mathematics. In whole class discussions Ms. Vance made a point of calling on girls and trying to encourage them to take part in discussions. This met with limited success. Ms. Vance took part in a series of workshops on gender issues in teaching, and one of the requirements was that she and a peer take turns observing each others' classes and recording the number of interactions that occurred between the teacher and each subgroup of students (by gender and ethnicity). The results of one of these observations (carried out by a fellow teacher) is given in Table 1.

Insert Table 1 about here

Two observations can be made about this data. First, it is evident that Ms. Vance was attempting to call on the girls at least as often as the boys. Second, the girls were less likely to respond than the boys.

Much like the other girls in the class, Janice, a tall and quiet girl, seldom took part in whole class discussions. In spite of Ms. Vance's cajoling, responses from Janice were likely to be monosyllabic, and she always avoided being asked to come up and demonstrate or give an explanation. She maintained this behavior throughout the fall, but in December, when Ms. Vance introduced the tape recorder to the class, a small change began to take place in Janice. Working with a partner and being interviewed by Ms. Vance, Janice enjoyed giving explanations to the microphone in the tape recorder. Her face would light up when such opportunities arose, and she would speak confidently and fluently about her work.

Over the next months, Ms. Vance used the tape recorder whenever students were working in pairs or small groups on an extended task. Janice, like many of her classmates, never hesitated to talk to the tape recorder when prompted. She continued to be reluctant to speak up in front of the whole class, though over time the confidence that she gained through the taped interviews began to affect her behavior with the whole class.

One day in mid-May I was observing Janice's class, along with Kathleen Heid, one of the evaluators for the Teacher Enhancement Project. The students were accustomed by now to my presence in the classroom, and usually took little notice. Dr. Heid was a new face, however, and she was introduced at the beginning of the class as a professor from Penn State who was an expert on algebra.

The lesson that day was on second degree equations. Ms. Vance was leading a whole class exploration with the TI-81 calculators on the effects of different values of A and C on the graphs of equations of the form Y = AX^2 + C. They had been working in pairs on an activity from the NCTM Addenda series on transforming graphs, but Ms. Vance stopped the work when, as she said, she witnessed lots of confusion and "poor mathematical language" being used. Mike had offered the conjecture that "if the number in front of x^2 is less than 1 it will make the parabola fatter, and if it's greater than 1 it will make it skinnier." After Tara had dictated this conjecture to Ms. Vance, who wrote it on the board, Janice raised her hand. She pointed out that "if we use a number smaller than 1, negative 1 is smaller than 1 but that turns it upside down." The class agreed as a whole to reword the conjecture. Ms. Vance then asked Janice what to do

BEST COPY AVAILABLE

about negatives. With a little coaxing, Janice was able to articulate that the negatives cause mirror images, but do not affect shrinking or stretching. Ms. Vance asked the class to write down their own conjecture about this, and Janice was the first to volunteer: "If the coefficient is negative, it doesn't change the shape, but turns it upside down in relation to the X axis." The discussion proceeded from there to horizontal shifts. When Ms. Vance asked what would move the graph horizontally, someone called out, "Add a number, like 3." Janice immediately put up her hand and said, "Wait, this is gonna go left 3, not right." Carl corrected her, and the discussion went on.

Though she was incorrect in her last point, it was quite clear that Janice had "found her voice" in the classroom. The typically silent student had freely joined the whole class discussion, offering to put her own words on the board for others to critique. After class Ms. Vance could not contain her pleasure. She expressed amazement that Janice had chosen this day, with two visitors to the class, to find the courage to do what she had been unwilling to do for nine months.

The case of Janice illustrates one of the unintended consequences of this experiment in observational assessment. When Ms. Vance chose to use a tape recorder as a preferred method of collecting evidence of student learning, she unintentionally helped Janice find a vehicle for expressing herself in the class. Over time, the confidence she had in her own mathematical thinking was enhanced to the extent that she was willing to express her mathematical opinion publicly to the whole class and even to outsiders.

Conclusions

The importance of this experiment for Ms. Vance and her students may lie more in its unexpected consequences than in any that were hypothesized from the start. I did not expect that the most critical component of the experiment would be in Ms. Vance's struggle to decide what to observe, nor did I anticipate the effect of the interviewing on students like Janice. In the NCTM *Assessment Standards for School Mathematics* (1995) assessment is described as an activity that occurs at the intersection of curriculum, instruction, and learning. These interconnections were apparent in Ms. Vance's experiment with an alternative means of assessment, in that the pivotal moment for making the assessment work came when she realized the importance of identifying her instructional goals. That process of creating a master rubric, in turn, brought a greater focus to her teaching and to the curricular choices she made. In other words, the assessment activities could not be separated from issues of instruction and curriculum, and all of these had some impact on student learning.

The integration of assessment with instruction, curriculum, and learning is also evident in the consequences of this experiment. The interaction between the teacher and the students, as shown by the types of questions Ms. Vance asked, was influenced by the assessment activities. When Ms. Vance was focused on recording what her students were doing and saying she tended to ask far more questions related to their strategies and explanations than when she was conducting a whole class discussion. A connection with student learning was seen in the case of Janice, who was emboldened by the small group interviews to participate in whole class discussions.

In some ways, this experiment was an investigation into the feasibility of putting

BEST COPY AVAILABLE

theory into practice. The theory of instructional assessment, calling for teachers to engage in interpretive research, was put to the practical test of daily life in the classroom. What we discovered was a number a limitations in the extent to which teacher can play this role. First, the opportunities for documenting observations and interviews were limited to the type of instruction on a given day. It was not feasible for Ms. Vance to collect this kind of data except when students were at the initial stages of work on an extended task. Second, she had to choose a method that was practical and required little time to use (in her case, the tape recorder). Third, the critical factor in collecting data was knowing precisely what she was looking for. It was not until she developed her master rubric that the documentation process became coherent and useful for her.

Unlike researchers, teachers do not have the time to pour over their data and look for patterns or trends. They are also severely limited, sometimes, in the kinds of interpretations they are required to make for the sake of giving evidence of student achievement Ms. Vance managed to solve these problems to a certain extent by developing the chart for tallying the problem solving progress of her students as she listened to the tape. Issues of analysis became issues of practicality. It was more difficult for her to find a way to aggregate this data with the other records she kept of her students' work, and to quantify the results.

From a practical point of view, this study gives some clues about two facets of classroom assessment. First, the importance of being clear about what is being assessed, and the links between that and instruction. Second, it highlights the practical challenges

of aggregating qualitative evidence with other sources of data for the purposes of monitoring student progress or assigning grades. There is a great need to clarify a theory of assessment that can take both of these facets into account and help teachers put the theory into practice.

The most important question resulting from this case study might be, Is observational assessment worth the time and energy it demands? This limited study gave only small indications of the possible positive consequences of a teacher's attempts to document student learning as she observed it in the classroom. Ms. Vance claimed that it brought greater focus to her teaching, and one student, Janice, seemed to benefit from the techniques she used. But more research is needed on the activities of teachers as researchers to determine the ultimate consequences for student learning.

References

Cobb, P. (1995, October). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. In D.Owens, M. Reed, & G. Millsaps (Eds.), *Proceedings of the seventeenth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, vol.1*. Columbus, O: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Dorr-Bremme, D.W. & Herman, J.L. (1986). *Assessing student achievement: A profile of classroom practices*. Los Angeles, CA: University of California, Center for the Study of Evaluation.

Graue, M.E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment, 1*(4), 283-309.

Ginsburg, H., Jacobs, S., & Lopez, L. (1993). Assessing mathematical thinking and learning potential in primary grade children. In M. Niss (Ed.), *Investigations into assessment in mathematics education*, pp.157-167. Netherlands: Kluwer.

Mellin-Olsen, S. (1993). A critical view of assessment in mathematics education: Where is the student as a subject? In M. Niss (Ed.), *Investigations into assessment in mathematics education*, pp.157-167. Netherlands: Kluwer.

Miller, L. (1991). Constructing pedagogical content knowledge from students' writing in secondary mathematics. *Mathematics Education Research Journal*, 3(1), 30-44.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

Schulman, (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Stenmark, J. (1991). *Mathematics assessment: Myths, models, good questions, and practical suggestions*. Reston, VA: National Council of Teachers of Mathematics.

van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Utrecht: Freudenthal Institute.

Vincent, M.L. & Wilson, L. (1996) Informal assessment: A story from the classroom. *Mathematics Teacher, 89* (3), 248-252.

Webb, N. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*, pp 661-686. New York: Macmillan.

Table 1

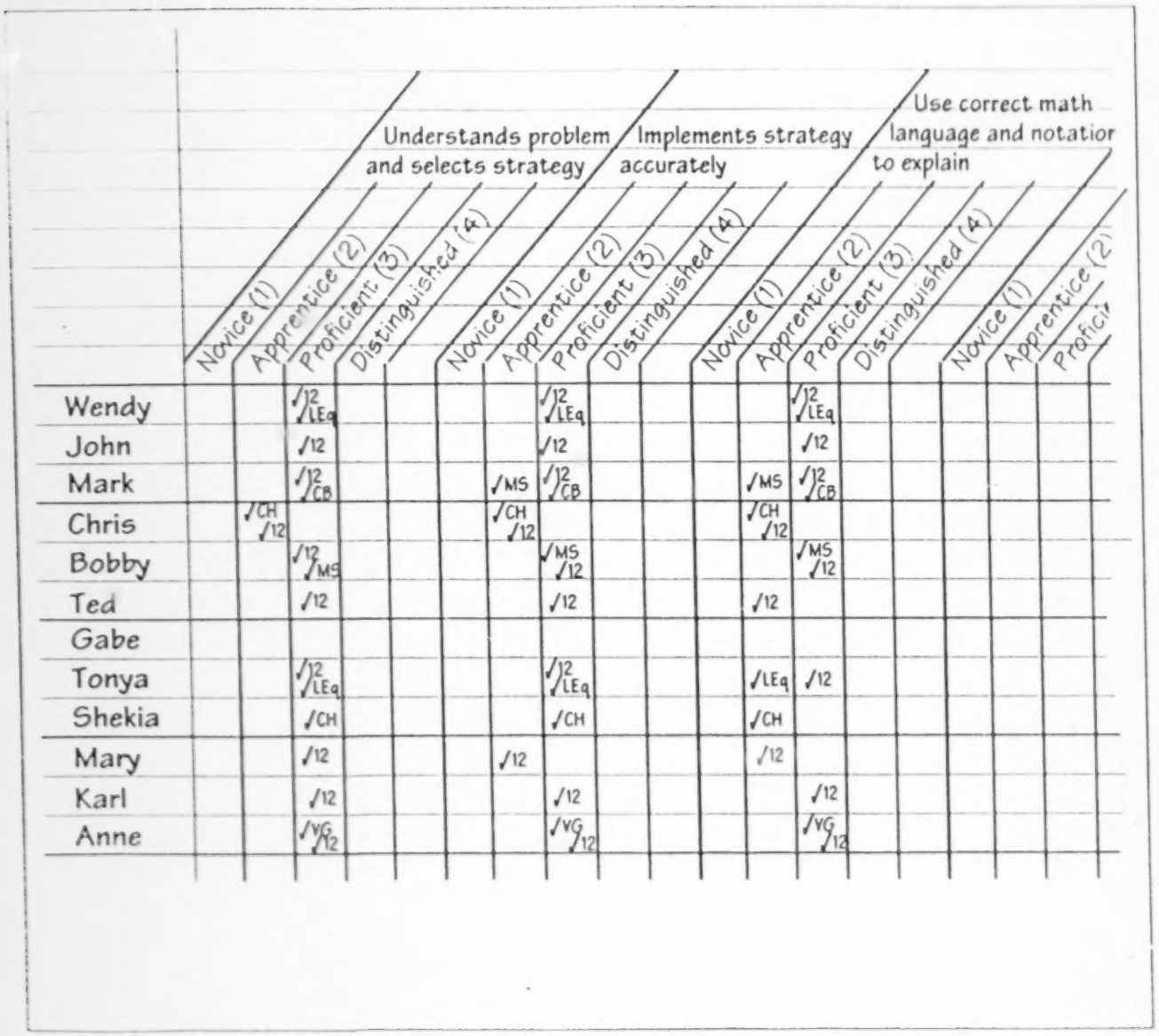
Responses by Gender and Ethnicity from One Class in October

Students	# in class	Calls on student	Calls on student	Response from student	Response from student
	(16 male, 10 female)	Male	Female	Male	Female
European-Americans	23	16	28	15	10
African-Americans	3	8	--	10	--

Figure Captions

Figure 1. Section of master rubric, used to record observations. The letters and numbers next to the check marks refer to different tasks.

Figure 2. Mark's work on the blackboard.


Understands problem and selects strategy
Implements strategy accurately
Use correct math language and notation to explain
Novice (1)
Apprentice (2)
Proficient (3)
Distinguished (4)
Wendy
John
Mark
Chris
Bobby
Ted
Gabe
Tonya
Shekia
Mary
Karl
Anne

$$1 + \underset{\times 3}{\underline{2}}\ = 7 + \underset{\times 5}{\underline{3}}\ = 22 + \underset{\times 7}{\underline{4}}\ = 50 + \underset{\times 9}{\underline{5}}\ = 95$$
$$\underline{6} \qquad \underline{15} \qquad \underline{28} \qquad \underline{45}$$

Appendix A

Ms. Vance's Master Rubric

UNDERSTANDS PROBLEM AND SELECTS STRATEGY

NOVICE

Indicates a basic understanding of problem and uses strategies

APPRENTICE

Indicates an understanding of problem and selects appropriate strategies

PROFICIENT

Indicates a broad understanding of problems with alternate strategies

DISTINGUISHED

Indicates a comprehensive understanding of problems with efficient, sophisticated strategies

IMPLEMENTS STRATEGY ACCURATELY

NOVICE

Implements strategies with minor mathematical errors in the solution

APPRENTICE

Accurately implements strategies with solutions

PROFICIENT

Accurately and efficiently implements and analyzes strategies with correct solution

DISTINGUISHED

Accurately and efficiently implements and evaluates sophisticated strategies with correct solutions

USES CORRECT MATHEMATICAL LANGUAGE AND NOTATION

NOVICE

Uses appropriate mathematical language some of the time

APPRENTICE

Uses appropriate mathematical language

PROFICIENT

Uses precise and appropriate mathematical language

DISTINGUISHED

Uses sophisticated, precise, and appropriate mathematical language throughout

USES A WIDE VARIETY OF REPRESENTATIONS

NOVICE

Uses few mathematical representations

APPRENTICE

Uses a variety of mathematical representations accurately and appropriately

PROFICIENT

Uses a wide variety of mathematical representations accurately and appropriately, uses multiple representations sometimes

DISTINGUISHED

Uses a wide variety of mathematical representations accurately and appropriately, uses multiple representations and states the connections

GIVES CLEAR EXPLANATIONS AND USES MATHEMATICALLY LOGICAL ARGUMENTS
NOVICE
Uses mathematical reasoning
APPRENTICE
Uses appropriate mathematical reasoning
PROFICIENT
Uses perceptive mathematical reasoning
DISTINGUISHED
Uses perceptive creative, and complex mathematical reasoning

DESCRIBES GENERAL SOLUTIONS AND FORMULATES NEW PROBLEMS
NOVICE
Describes general solutions using words
APPRENTICE
Describes general solutions using algebraic notation
PROFICIENT
Describes general solutions algebraically and tests the solution on a new problem
DISTINGUISHED
Describes general solution algebraically, test the solution on new problem and make connections to other problems

WORKS COOPERATIVELY LISTENS AND CONTRIBUTES
NOVICE
Participates in group activities
APPRENTICE
Participates in group activities and considered the ideas of others
PROFICIENT
Participates in group activities, illicites group input and seriously considers the ideas of others
DISTINGUISHED
Participates in group activities and involves others by requesting input or challenging others

SHOWS WILLINGNESS TO TAKE RISKS AND PERSEVERE
NOVICE
Attempts difficult tasks
APPRENTICE
Shows willingness to tackle hard tasks and perseveres
PROFICIENT
Enjoys solving problems and perseveres
DISTINGUISHED
Shows confidence in problem solving and perseveres until a solution is acceptable

CORRECTS , EDITS AND REVISES
NOVICE
Recognizes errors
APPRENTICE
Shows evidence of self-assessment and self-correction
PROFICIENT]
Consistently self-assesses and self-corrects
DISTINGUISHED
Consistently self-assesses and self-corrects and is able to reflect on own growth